Equilibrium Outcomes in Repeated Two-Person, Zero-Sum Games

Guilherme Carmona
Universidade Nova de Lisboa

October 15, 2002

Abstract

We will consider repeated two-person, zero-sum games in which the preferences in the repeated game depend on the stage-game preferences, although not necessarily in a time-consistent way. We will assume that each player's repeated game payoff function at each period of time is strictly increasing on the stage game payoffs and that the repeated game is itself a zero-sum game in every period. Under these assumptions, we will show that an outcome is a subgame perfect outcome if and only if its components are all Nash equilibria of the stage game.

1 Introduction

Recently, there has been an increasing interest in time-inconsistent, and its consequences for economic theory, and policy. Much of this interest was motivated by the work of Laibson [4], Jovanovic and Stolyarov [2], and Kocherlakota [3], which have shown that time-inconsistent preferences can change a model's implications for economic policy. In contrast, we will show that in a two-person, zero-sum game time-inconsistent preferences have no effect over the equilibrium outcomes that can arise, and so have no effect over the model's implication for economic policy.

Our main result can be stated as follows: consider a repeated two-person, zero-sum games in which the preferences in the repeated game depend on the stage-game preferences, although not necessarily in a time-consistent way.

Assume that each player's repeated game payoff function at each period of time is strictly increasing on the stage game payoffs, and, if the repeated game is itself a zero-sum game in every period. Under these assumptions, we will show that an outcome is a subgame perfect outcome if and only if all its components are all Nash equilibria of the stage game.

Given the recent interest on the economic effects of time-inconsistent preferences, it is interesting to know what game-theoretic results change by assuming time-inconsistent preferences. This question seems natural to us since, as Peleg and Yaari [5] and Goldman [1] pointed out, the appropriate way of modelling time-inconsistency in preferences is through the concept of subgame perfect equilibrium of a game between an agent and his future selves.[1] Therefore, it seems necessary to know how game-theoretic results will change in order to understand the economic effects of time-inconsistent preferences.

It should be noted that the conclusion of our main result belongs to the oral tradition of game theory, at least when the repeated game payoffs are given by the discounted sum of stage game payoffs. The contribution of our work is to provide a simple proof of that result, and to show that it holds under quite general assumptions.

2 Notation and definitions

A *two-person, zero-sum game* G is defined by

$$G = (A_1, A_2, u_1, u_2)\,,$$

where for all $i = 1, 2$: (1) A_i is a finite set of player i' actions, and (2) $u_i : A \to \mathbb{R}$, where $A = A_1 \times A_2$, is player i' payoff function; the player's payoff functions satisfy

$$u_1(a) + u_2(a) = 0,$$

for all $a \in A$. Let $S_i = \Delta(A_i)$, $S = S_1 \times S_2$, and $u_i : S \to \mathbb{R}$ be the usual extension to mixed strategies.

Let, for $i = 1, 2$, $v_i = \min_{s_{-i}} \max_{s_i} u_i(s_i, s_{-i})$, and $NE = \{s \in S : \text{for all } i = 1, 2, u_i(s) \geq u_i(\tilde{s}_i, s_{-i}), \text{ for all } \tilde{s}_i \in S_i\}$. The set NE is the set of Nash equilibria of G, and v_i is the minmax level for player i.

[1] The concept of time-inconsistent preferences was itself introduced by Strotz [6].

The *supergame of* G consists of an infinite sequence of repetitions of G taking place in periods $t = 1, 2, 3, \ldots$. At period t the players make simultaneous moves denoted by $s_i^t \in S_i$ and then each player learn his opponent's move.

For $k \geq 1$, a $k-$*stage history* is a $k-$length sequence $h_k = (s^1, \ldots, s^k)$, where, for all $1 \leq t \leq k$, $s^t \in S$; the space of all $k-$stage histories is H_k, i.e., $H_k = S^k$ (the $k-$fold Cartesian product of S.) The notation e stands for the unique 0–stage history — it is a 0–length history that represents the beginning of the supergame. The set of all histories is defined by $H = \bigcup_{n=0}^{\infty} H_n$.

For every $h \in H$, define $h^r \in S$ to be the projection of h onto its r^{th} coordinate. For every $h \in H$ we let $\ell(h)$ denote the *length of* h. For two positive length histories h and $\bar{h}$ in H we define the *concatenation of* h *and* $\bar{h}$, in that order, to be the history $(h \cdot \bar{h})$ of length $\ell(h) + \ell(\bar{h})$: $(h \cdot \bar{h}) = (h^1, h^2, \ldots, h^{\ell(h)}, \bar{h}^1, \bar{h}^2, \ldots, \bar{h}^{\ell(\bar{h})})$. We also make the convention that $e \cdot h = h \cdot e = h$ for every $h \in H$.

It is assumed that at stage k each player knows h_k, that is each player knows the actions that were played in all previous stages. Regarding strategies, players chose behavioral strategies, that is, in each stage k, they choose a function from H_{k-1} to S_i denoted f_k^i, for player $i = 1, 2$. The set of player i's strategies is denoted by F_i, and $F = F_1 \times F_2$ is the joint strategy space. Finally, a strategy vector is $f = \left(\{f_k^i\}_{k=1}^{\infty}\right)_{i=1,2}$.

Given an individual strategy $f_i \in \Sigma_i$ and a history $h \in H$ we denote the *individual strategy induced by* f_i *at* h by $f_i|h$. This strategy is defined pointwise on H: $(f_i|h)(\bar{h}) = f_i(h \cdot \bar{h})$, for every $\bar{h} \in H$. We will use $(f|h)$ to denote $(f_1|h, \ldots, f_n|h)$ for every $f \in S$ and $h \in H$.

Any strategy $f \in F$ induces an outcome $\pi(f)$ as follows:

$$\pi^1(f) = f(e) \quad \pi^k(f) = f(\pi^1(f), \ldots, \pi^{k-1}(f)),$$

for $k \in \mathbb{N}$. Thus, we have define a function $\pi : F \to S^{\infty}$, where $S^{\infty} = S \times S \times \cdots$.

Let $M \geq 0$ be such that $|u_i(s)| \leq M$, for all $s \in S$, and $i \in N$. Then, any outcome $\pi \in S^{\infty}$ induces two elements in l^{∞}, one for each player, as follows

$$x_i^k(\pi) = u_i(\pi^k),$$

for all $k \in \mathbb{N}$. Thus, we have define a function $x_i : S^{\infty} \to l^{\infty}$, for all $i = 1, 2$.

For $x, y \in l^\infty$, $x \geqq y$, means $x_k \geq y_k$, for all $k \in \mathbb{N}$; $x \geq y$ means $x \neq y$ and $x \geqq y$.

Let for each $i = 1, 2$, and $k \in \mathbb{N}$, $U_i^k : l^\infty \to \mathbb{R}$ be given. The payoff for player i, $i = 1, 2$, from his point of view in period $k \in \mathbb{N}$ of a strategy $f \in F$ in the supergame of G is defined to be $U_i^k(x_i \circ \pi(f))$.

A strategy vector $f \in F$ is a *subgame perfect equilibrium* of the supergame of G if $U_i^k(x_i(h \cdot \pi(f|h))) \geq U_i^k(x_i(h \cdot \pi(g_i, f_{-i}|h)))$, for all $i = 1, 2$, $k \in \mathbb{N}$, $h \in H_{k-1}$ and $g_i \in F_i$. Let $E\Pi$ denote the set of subgame perfect equilibrium outcomes.

3 Equilibrium outcomes

In this section we state and prove our main result.

Theorem 1 *Suppose that for all $k \in \mathbb{N}$, and $i = 1, 2$,*

1. $U_1^k(x_1(\pi)) + U_2^k(x_2(\pi)) = 0$, *for all* $\pi \in S^\infty$,
2. U_i^k *is strictly increasing:* $x, y \in l^\infty$ *and* $x \geq y$ *implies* $U_i^k(x) > U_i^k(y)$.

Then, $E\Pi = NE^\infty$ *and* $u_i(\pi^k) = v_i$ *for all* $\pi \in E\Pi$, $i = 1, 2$, *and* $k \in \mathbb{N}$.

Proof. Let $\pi \in E\Pi$, $i = 1, 2$, and $k \in \mathbb{N}$. By 2,

$$U_i^k(x_i(\pi)) \geq U_i^k(x_i^1(\pi), \ldots, x_i^{k-1}(\pi), \max_{s_i} u_i(s_i, \pi_{-i}^k), v_i, v_i \ldots) \geq$$
$$\geq U_i^k(x_i^1(\pi), \ldots, x_i^{k-1}(\pi), v_i, v_i, \ldots) := \bar{v}_i^k.$$

Let α be a Nash equilibrium of G; thus, in particular, $u_i(\alpha) = v_i$. By letting $\tilde{\pi} = (\pi^1, \ldots, \pi^{k-1}, \alpha, \alpha, \ldots)$, we conclude by 1 that $\bar{v}_1^k + \bar{v}_2^k = U_1^k(x_1(\tilde{\pi})) + U_2^k(x_2(\tilde{\pi})) = 0$. Also, by 1, $U_1^k(x_1(\pi)) + U_2^k(x_2(\pi)) = 0$. Hence, $U_k(x_i) = \bar{v}_i^k$.

We therefore conclude that,

$$U_k(x_i^1(\pi), \ldots, x_i^{k-1}(\pi), \max_{s_i} u_i(s_i, \pi_{-i}^k), v_i, v_i \ldots) =$$
$$= U_k(x_i^1(\pi), \ldots, x_i^{k-1}(\pi), v_i, v_i, \ldots),$$

and so by 2, $\max_{s_i} u_i(s_i, \pi_{-i}^k) = v_i$.

Since $u_i(\pi^k) = x_i^k(\pi) \leq \max_{s_i} u_i(s_i, \pi_{-i}^k) = v_i$, for all $k \in \mathbb{N}$, and $U_i^1(x_i(\pi)) \geq U_i^1(v_i, v_i, \ldots)$, it follows that

$$u_i(\pi^k) = v_i = \max_{s_i} u_i(s_i, \pi_{-i}^k);$$

hence, π^k is a Nash equilibrium. ■

Corollary 1 *Suppose that for all $k \in \mathbb{N}$*

1. *$U_i^k = U_k$, for $i = 1, 2$,*
2. *U_k is additive: $U_k(x + y) = U_k(x) + U_k(y)$, for all $x, y \in l^\infty$,*
3. *U_k is strictly increasing: $x, y \in l^\infty$ and $x \geq y$ implies $U_k(x) > U_k(y)$.*

Then, $E\Pi = NE^\infty$ and $u_i(\pi^k) = v_i$ for all $\pi \in E\Pi$, $i \in N$, and $k \in \mathbb{N}$.

The following example shows that we cannot dispense with additivity. The stage game is the matching pennies:

$1\backslash 2$	H	T
H	$1, -1$	$-1, 1$
T	$-1, 1$	$1, -1$

Assume time-consistency, and let $w = (-1, 1, -1, 1, ...)$.
The preferences are, for $\delta \in (0, 1)$ and $M > 0$,

$$U(x) = \begin{cases} (1-\delta)\sum_{k=1}^{\infty} \delta^{k-1} x_k + M & \text{if} \quad x \geqq w \\ (1-\delta)\sum_{k=1}^{\infty} \delta^{k-1} x_k & \text{otherwise.} \end{cases}$$

This preferences are strictly increasing because $\geqq$ is transitive and the discounted sum is strictly increasing. Define π as follows $\pi^1 = (H, H)$, $\pi^2 = (T, H)$, $\pi^3 = (H, H)$, $\pi^4 = (T, H), \ldots$, and define

$$f_i(h) = \begin{cases} \pi_i^k & \text{if} \quad h = (\pi^1, ..., \pi^{k-1}) \\ \text{play H with 1/2 probability otherwise.} \end{cases}$$

Then f is a subgame perfect equilibrium and π is a SPE outcome (and it doesn't consist of Nash equilibria of the stage game). This is so because the payoff of the equilibrium path is $M - \frac{(1-\delta)^2}{1-\delta^2}$, and the payoff from deviating is 1. By choosing M big enough, we can deter deviations.

Let for $h \in H$ and $r \in \mathbb{N}$, $\lambda^r h := (h^1, \ldots, h^r)$ and $\mu^r h := (h^r, h^{r+1}, \ldots)$. We say that U_k, $k \in \mathbb{N}$, is *independence of the past* if for all $x, y \in S^\infty$ satisfying $\lambda^{k-1}x = \lambda^{k-1}y$ then $U_k(x) \geq U_k(y)$ if and only if $U_k(\mu^k x) \geq U_k(\mu^k y)$. If, for all $k \in \mathbb{N}$, U_k is independent of the past, then $f \in F$ is a subgame perfect equilibrium if and only if $U_k(x_i \circ \pi(f|h)) \geq U_k(x_i \circ \pi(g_i, f_{-i}|h))$, for all $i = 1, 2$, $k \in \mathbb{N}$, $h \in H_{k-1}$ and $g_i \in F_i$.

Theorem 2 *Suppose that for all $k \in \mathbb{N}$*

1. *$U_i^k = U_k$, for $i = 1, 2$,*

2. *U_k is independent of the past,*

3. *$U_k(x + y) \geq U_k(x) + U_k(y)$, for all $x, y \in l^\infty$,*

4. *$U_k(\alpha, \alpha, \ldots) = \alpha$, for all $\alpha \in \mathbb{R}$,*

5. *U_k is strictly increasing: $x, y \in l^\infty$ and $x \geq y$ implies $U_k(x) > U_k(y)$.*

Then, $E\Pi = NE^\infty$ and $u_i(\pi^k) = v_i$ for all $\pi \in E\Pi$, $i = 1, 2$, and $k \in \mathbb{N}$.

Proof. Let $\pi \in E\Pi$, $i = 1, 2$, and $k \in \mathbb{N}$. For $t \geq k$, let $x_i^t = u_i(\pi^t)$, and $x_i = (x_i^k, x_i^{k+1}, \ldots)$. By 3, and 4,

$$U_k(x_i) \geq U_k(\max_{s_i} u_i(s_i, \pi_{-i}^k), v_i, v_i \ldots) \geq U_k(v_i, v_i, \ldots) = v_i.$$

By 2, and 3,

$$U_k(x_1) + U_k(x_2) \leq U_k(x_1 + x_2) = 0.$$

Because $v_1 + v_2 = 0$, it follows that

$$U_k(x_i) = v_i.$$

Thus, by 4, $\max_{s_i} u_i(s_i, \pi_{-i}^k) = v_i$. Since $u_i(\pi^k) = x_i^k \leq \max_{s_i} u_i(s_i, \pi_{-i}^k) = v_i$, then $u_i(\pi^k) = v_i$. Since this equality holds for all $k \in \mathbb{N}$, it follows that $x_i = (v_i, v_i, \ldots)$, and so by 4,

$$u_i(\pi^k) = v_i = \max_{s_i} u_i(s_i, \pi_{-i}^k);$$

hence, π^k is a Nash equilibrium. ■

References

[1] Goldman, S., (1980), "Consistent Plans", *Review of Economic Studies*, **47**, 533-537.

[2] Jovanovic, B., and D. Stolyarov, (2000), "Ignorance is Bliss," Manuscript, New York University.

[3] Kocherlakota, N., (2001), "Looking for Evidence of Time-Inconsistent Preferences in Asset Market Data," *Federal Reserve Bank of Minneapolis Quarterly Review*, **v. 25, no. 3**, 13-24.

[4] Laibson, D., (1997), "Golden Eggs and Hyperbolic Discounting," *Quarterly Journal of Economics*, **112**, 443-477.

[5] Peleg, B., and M. Yaari (1973), "On the Existence of a Consistent Course of Action when Tastes are Changing," *Review of Economic Studies*, **40**, 1-24.

[6] Strotz, R. (1955), "Myopia and Inconsistency in Dynamic Utility Maximization," *Review of Economic Studies*, **23**, 165-180.